My goal in this chapter is to lay out some claims about the initial cognitive structures that underlie our earliest understanding of social entities and their interactions, and that serve as the foundations upon which more elaborated social understanding is built. I propose three innately given components of understanding as our entrée to understanding the social world:

First: The social and physical worlds are viewed, from early infancy, as fundamentally distinct, and separate cognitive systems are engaged in reasoning about the two. Infants reason about inanimate objects in terms of naïve physics, while they reason about social beings in terms of mental attributes such as goals, desires, and attitudes. These systems are sufficiently distinct that under at least some conditions, infants do not even think of social beings as material objects at all. Indeed, infants view the actions of intentional agents as ontologically different from the actions of inanimate objects. They make “essentialist” inferences about the source of behavior of intentional agents, viewing their actions as arising from intrinsic, essential features, not superficial or extrinsic attributes.

Second: The goals, desires, and attitudes that young infants ascribe to social beings, and via which they interpret their behavior, are inherently mentalistic attributions. Many studies show that infants as young as three to six months can attribute a goal to a social agent; but researchers take different positions on whether the infant’s “goal” concept is properly viewed as a genuinely mentalistic one or not. On the basis of research findings from our lab, I suggest that infants’ earliest understanding of social entities and their interactions involves reference to the mental contents and states of the actors.

Third: Infants are not interpreting interactions among agents and assigning attitudes dispassionately, but are judging these social interactions. The actions of one intentional being toward another are viewed in valenced terms, as “good” or “bad,” from early infancy; they are understood and assessed relative to their social effects, in terms of their positive and negative impacts upon (the goals and desires of) other social beings. Moreover, these judgments are not restricted to actions themselves but are extended to the actors—individuals are evaluated in terms of their social behavior. These evalua-
tions undergird our very earliest analyses of the social world, and may serve as the initial foundation upon which a system of moral cognition is subsequently built.

1 Reasoning About Intentional Agents Versus Inanimate Objects

Since roughly the mid-1980s, researchers have been extensively exploring what infants know about intentional agents, and whether infants distinguish agents from other kinds of things in the world. Results from many studies suggest that infants treat intentional agents—or, at least, the specific subcategory of human agents—differently from other entities, even in the first days and weeks of life. Newborn infants prefer to look at faces (particularly those with open eyes; Batki et al., 2000) over other stimuli (e.g., Slater and Quinn, 2001; M. H. Johnson, 2001). Infants also react differently to the behavior of intentional agents than to that of other objects. When three-month-old infants see an inanimate moving object (such as a rolling ball or stone) cease moving and become motionless, they lose interest. But infants become unhappy upon seeing a responsive human face stop moving and become nonresponsive (the classic “Still Face” paradigm; Corkum and Moore, 1998; Tronick et al., 1978). Young infants also expect people, but not other objects, to respond contingently to their—the infants’—own actions. Two-month-olds greatly enjoy interacting with their mother via closed-circuit TV in real time, but become distressed if her responses are presented with even brief a temporal lag, as this obscures the contingent nature of her responses to the infant’s own actions and responses (Murray and Trevarthen, 1985). And nine-month-olds show a preference for moving dot patterns that correspond to a walking human, as opposed to other equally complex but nonbiological motion patterns (Bertenthal et al., 1985). These and many other studies show that infants discriminate social agents from inanimate entities, understand something of their typical behavior, and are very interested in and oriented toward the social world.

1.1 The Physics of Agents

In my lab, we have been examining what infants understand about the behavior of social entities and, in particular, asking whether infants’ system for reasoning about the social world is distinct from their reasoning about the physical world. Do infants appreciate that social entities are different in interesting and important ways from inanimate material entities? To answer this first requires knowing how infants reason about the material world. An extensive body of research since the mid-1980s by Elizabeth Spelke (e.g, Spelke, 1990, 1994) has elucidated principles that govern infants’ expectations about inanimate material objects. Infants as young as three months of age expect objects

(1) to be spatiotemporally continuous—to follow a continuous pathway through space and to have an enduring existence through time;
(2) to be solid—such that two distinct objects cannot occupy the same space at the same time, and one object cannot pass through another.
(3) to be cohesive—maintaining their distinct boundaries through their interactions with each other, not spontaneously merging with other objects or disassembling into separate objects;

(4) to influence each other only through direct physical contact—such that one object cannot influence another at a distance.

These principles provide the basis for infants’ expectations about events involving material objects and how these events will unfold. And at least some of these principles—the cohesion principle, for example—serve as the defining criteria for identifying material objects in the world: A thing which violates certain of these principles is not considered an object. Sand, for example, is not cohesive—it pours, separates into distinct portions, and recombines in a fluid fashion in which the boundaries of an individual portion are not retained and preserved. And infants, when shown a portion of sand being poured, do not construe the portion of sand as an object—they do not apply expectations of spatiotemporal continuity to the portion (Huntley-Fenner, Carey, and Solimando, 2002). This is true not only for sand, but also for other portions that behave in a noncohesive manner, such as a jumbled-together pile of small Lego blocks (Chiang and Wynn, 2000). That is, something that violates cohesion is not viewed as an object, and so the remaining object principles are not applied.

Intentional agents, such as people, are an interesting kind of material object. Like inanimate objects, we are solid, cohesive, material beings that take up space and move continuously through space and time. However, intentional beings are not constrained by the principle of contact—we can influence each other from a distance, and we need not be physically contacted by another object or person in order to initiate movement. This raises the question of whether infants understand the ways in which intentional agents are, and are not, constrained by the principles that constrain inanimate object behavior. Spelke, Phillips, and Woodward (1995) explored whether infants extend the principle of contact to people, and found that, in contrast to their reasoning about physical objects, infants do not think that people must first be contacted in order to move on their own (see also Poulin-Dubois et al., 1996; Rakison and Poulin-Dubois, 2001, for related findings in infants). That is, infants seem to reason about an agent’s movements in much the same way as do adult humans, and to appreciate that agents initiate motion in different ways than non-agents.

These findings are consistent with the position that infants have a mature understanding of the similarities and differences between social agents and inanimate objects. My colleagues and I, however, propose a very different position: Contrary to setting up appropriate expectations about the physical nature of intentional agents, infants reason about agents in a very different way than do adults. Categorizing something as an intentional agent may impede the application of physical object principles, perhaps because a more salient set of social principles has been activated, or perhaps because intentional agents violate the principle of contact, which might lead infants to conclude, erroneously, that they are not physical objects (as when infants see a portion of sand violate cohesion, and conclude—correctly, in this case—that it is not an object). We suggest that the system of object understanding
and the system of agent understanding are mutually exclusive in infants, such that if a thing is construed as an agent (with the corresponding set of agent principles activated), that thing is not thought of as a material object, and vice versa (Bloom, 2004; Kuhlmeier et al., 2004).

To test this possibility, we ran two experiments to explore whether five-month-old infants apply the expectation of spatiotemporal continuity both to people and to inanimate blocks, using a modified version of a methodology previously employed by Spelke and colleagues (Spelke and Kestenbaum, 1986; Spelke, Kestenbaum, Simons, and Wein, 1995). In our first experiment, a basic replication of Spelke and colleagues’ study, two groups of infant subjects were habituated to videotapes of inanimate objects—in our study, white and blue boxes—following either a continuous or a discontinuous path of motion (Kuhlmeier, Bloom, and Wynn, 2004; see also figure 17.1a).

In the continuous motion event, infants were shown a sequence that adult observers perceive as a single object moving back and forth behind two occluding
barriers: A box moved from the left side of the stage and behind the left barrier. An identical box was then seen moving rightward, appearing from behind the left barrier and moving toward, and then behind, the right barrier. An identical box then moved out from behind the right barrier, onto the right side of the stage, and then back behind the right barrier, shortly followed by a box appearing from behind the right barrier, headed leftward toward, and behind, the left barrier, then out from the left side of the left barrier. This entire sequence looped repeatedly on each trial, until the infant looked away, signaling the end of the trial.

The discontinuous motion event was identical, except that no object ever traversed the space between the two barriers; adult observers perceive this sequence as an event containing two different objects moving in the display. After habituation, subjects were presented with two test events: a one-box event, in which a single object moved continuously back and forth across the full width of the stage, or a two-box event, in which two objects moved across the stage.

In our second experiment, we presented two further groups of infants with videotapes of people instead of boxes, following the same motion patterns— a person walked in a continuous motion pathway behind the screens for one group of infants, and in a discontinuous motion pathway for the other group of infants (see figure 17.1b).

Replicating the results of Spelke and colleagues, the infant subjects in our boxes experiment applied expectations of continuity to the inanimate objects: Infants in the continuous and discontinuous box conditions showed reliably different patterns of looking at the two test events, the former looking longer at the two-box test event and the latter looking longer at the one-box test event.

But infants did not use this principle to reason about the motions of people; infants in both the continuous and discontinuous persons conditions looked equally at one-person and two-persons test events. Infants did not use the path of motion (continuous or discontinuous) of the observed person to discern whether one or two distinct individuals were present in the scene. This suggests that five-month-olds (mistakenly) do not apply the physical constraint of continuity to human agents. That is, infants at five months of age are not viewing humans as material objects. Two possible explanations for this result occur to us.

First, this finding could reflect a relative salience issue: The material nature of inanimate objects may simply be more salient to infants, whereas they are so busy thinking about the other attributes of people and intentional agents that they simply neglect at times to apply the constraints of their naïve physics. A study by Saxe, Tzelnic, and Carey (2005) supports this interpretation: Their study found that five-month-old infants did not expect a human hand to be able to pass through a solid wall. Infants saw a wall, which was then occluded by a screen placed in front of it. Infants then saw an experimenter’s arm reach behind the screen and the experimenter’s hand appear on the other side of the screen, as if the arm were reaching straight through the solid wall. Infants gave longer looks to this event than to one in which there was no wall behind the screen. This suggests that infants do, in some circumstances, appreciate that humans are material objects constrained by the principle of solidity. However, it’s important to note that in the impossible event, part of the experimenter’s arm was, at that very moment, occupying some of the same space as the wall—to mentally connect the visible portion of arm and hand required representing the hidden portion of the arm in the process of passing through the wall.
It is possible that infants were responding not to the violation of solidity this event presents, but to the complete novelty of an event in which a person and an object are “overlapped” like this. The same infants might well fail to invoke the solidity principle to constrain their assumptions about the possible trajectories of humans. An experiment to test whether infants use the continuity principle to constrain their inferences about people’s possible paths of motion would be to ask whether infants use the presence of a solid wall between two locations, A and B, to infer that a person at A is unable to get to B. Shown a person at A, would infants exhibit surprise upon subsequently seeing that person at B? There is no empirical evidence to suggest what infants would do in such a test.

1. Perhaps if we ran the same experiment, but first “reminded” the infants that the people in the movies were in fact physical objects (by, for example, having them bang into a wall or table and say “Ouch!”), thus exhibiting adherence to the solidity principle, infants would then be prompted to apply the principle of continuity in the task and respond in the “people” version of the experiment as they do in the “boxes” version.
The second possibility is that, for the infant, being a material object is inconsistent with being an intentional agent. To be an agent entails belonging to a fully distinct and nonoverlapping category of entity, with its own distinct governing principles. This would entail that people are so different from physical objects in how infants think of them, that they are not even conceived of as physical objects at all. This claim fits in with other research indicating the difficulty in representing the dual natures of things (e.g., DeLoache, 2000; Pierroutsakos and DeLoache, 2003). If this is so, then the developmental task for infants is to learn how to integrate these two distinct systems of reasoning, for of course intentional agents are also material objects, and as adults we are able to reason about both the intentional and the physical aspects of people and other intentional agents.

1.2 The Insides of Agents

How do infants know to which entities to apply their system of reasoning about social entities; that is, how do infants identify those things in the world that are intentional beings? The evidence suggests that they use a range of different cues, any one of which may be sufficient to engage intentional attribution. These cues may include morphological features, such as the presence of a face (e.g., S. C. Johnson, Slaughter, and Carey 1998); indications of social responsiveness, such as contingently responding to another social entity (Arita, Hiraki, Kanda, and Ishiguro 2005; S. C. Johnson et al., 1998); indications of rationality, such as moving toward a goal in the most efficient manner (e.g., Gergely et al., 1995; Gergely and Csibra, 1997); and certain types of movement, such as self-generated motion that violates inertial forces (e.g., Premack and Premack, 1997; Kuhlmeier et al., 2003; Kuhlmeier, Wynn, Bloom, and DeCoste, 2005).

In one study, we asked whether motion cues alone would lead infants to view an entity as goal-driven (Kuhlmeier, Wynn, Bloom, and DeCoste, 2005). We adapted our methods from the now-classic study by Amanda Woodward, in which infants, habituated to a hand reaching repeatedly to one of two toys, expect the hand to continue to reach to that same toy when the locations of the toys are switched (they look longer when the hand reaches to the same old location which now houses the other toy), but do not hold these expectations of an inanimate object (e.g., Woodward, 1998). We habituated 14-month-old infants to a computer-animated scenario in which a little ball repeatedly approached a particular one of two objects (a square and a triangle) visible in the display. In our Animate-Motion condition (figure 17.2, top panel), the little ball showed self-generated, noninertial motion. It commenced its motion from a standstill, moved forward, wavered briefly, then oriented directly toward one of the objects and went straight to it, stopping when it reached the object. In our Inertial-Motion condition (figure 17.2, bottom panel), these motion cues to animacy were not present: The ball entered from offscreen, already rolling in a straight trajectory toward the object when it appeared on screen. Following habituation, we showed infants test trials in which the positions of the triangle and square were reversed. In one kind of test trial, the little ball approached the original object, now in its new location; in the other kind of test trial, the little ball approached the original location, which now held the object it had never approached during habituation. If infants viewed the little ball as goal-driven, they should respond as infants do in Woodward’s many studies. They should look longer when the ball follows the
same, habituated spatial pathway to the new object, than when the ball takes a new rolling pathway to reach the object it approached on the habituation trials.

This is the pattern of results we found in our Animate-Motion condition; infants looked significantly longer when the ball approached the new object, indicating that in the habituation trials, they had viewed the ball as being a goal-directed (i.e., intentional) agent, and were now responding to the ball’s apparent change in goal. But we obtained the reverse pattern in our Inertial-Motion condition: Infants looked significantly longer when the ball rolled to the same old object in the new location than when it rolled along the habituated spatial path. That is, infants habituated to a ball manifesting an inertial pattern of motion did not view the ball as a goal-directed, intentional agent. Thus, self-generated movement that violates inertial principles is, in itself, a sufficient cue for babies to view an entity as an intentional agent.

**Figure 17.2** The animate-motion and inertial-motion habituation and test events presented to infants in Kuhlmeier, Wynn, Bloom, and DeCoste (2005). Infants in the animate-motion event look longer on new goal/old path test trials, while infants in the inertial-motion event look longer on old goal/new path test trials.
Why is this so? It could be that the presence of self-generated motion, like the presence of certain morphological features (such as facial features), automatically activates notions of intentionality; that there is a built-in mapping between this feature per se and the concept of intentional agent. I propose a somewhat different possibility: that self-generated movement is not a feature that automatically cues intentionality per se, but that it cues agency precisely because it suggests the presence of a causal power internal to the entity.

Adults appreciate that the causal force behind an agent’s self-generated actions comes from inherent properties of the agent itself. It is not the product of some external force, as when a leaf is blown by the wind. In addition, it is typically seen as the outcome of an internal property of the agent—sometimes described as the agent’s essence (e.g., Bloom, 2000; S. Gelman, 2003). While there is much research exploring whether this assumption is held by young children (e.g., R. Gelman et al., 1995; Inagaki and Hatano, 2002; Newman, Cheries, Keil, and Bloom, under review; Simons and Keil, 1995), little work has been done with infants. In our next studies, we examined whether infants recognize that an agent’s ability to act in a given manner stems from intrinsic properties, as opposed to more transient aspects of its perceptual experience.

Our findings support this notion, suggesting that infants differentially weight select features of an object in their reasoning about the causal locus of the object’s actions, depending on whether the object is one that exhibits animate patterns of motion. Specifically, our findings suggest that infants’ inferences about the actions of entities that produce self-generated, noninertial motion reference deep, intrinsic properties of the entity, while their early inferences about the actions of inanimate objects reference superficial, nonessential properties (Newman, Herrmann, Wynn, and Keil, in press).

In our first of these experiments, we presented 14-month-old infants with two computer-animated, catlike creatures. Each cat exhibited its own type of animate motion—one swayed back and forth, and the other bounced up and down. Moreover, each cat was characterized by two features: one intrinsic to the animal (the animal’s abdomen) and one extrinsic to the creature (a hat it was wearing). These creatures were discriminable not only in the motion they made, but also in the colors of their features: One cat had a red abdomen and wore a red hat, while the other had a blue abdomen and wore a blue hat. Following habituation to these creatures, in which each cat repeatedly performed its unique movement, infants were presented with a new cat, which had the abdomen color of one cat and the hat color of the other (that is, a red abdomen and blue hat, or vice versa). We asked simply how infants expected the novel cat to move—like the cat who shared its abdomen color, or the cat who shared its hat color. That is, which feature would infants selectively weight as the relevant feature along which to generalize a kind of animate motion—a feature intrinsic to the animal making the motion (its abdomen color), or a feature belonging to a superficial and separable component (its hat color)? Our infants strongly prioritized the intrinsic feature over the extrinsic feature when predicting how the novel individual would move—they expected it to make the motion of the habituation cat that shared its abdomen color, and looked significantly longer when it made the motion of the other cat.

In a second experiment, we showed that this selective weighting of intrinsic features holds only for objects that undergo self-generated, noninertial motion (that is, precisely the kind of motion that signals intentional agency to infants). If an object undergoes this same motion but with an obvious external cause, infants do
not generalize on the basis of an internal feature. We presented 14-month-old infants with a real object (a translucent, animal-like toy) that exhibited a highly salient and attractive (to infants) wiggling/shaking motion. For one group of infants, the object made this motion autonomously (the Self-Generated Motion condition). For another group of infants, the object was caused to make this motion through external means (being shaken by the experimenter; the Externally Caused Motion condition). Importantly, for both groups of infants, the same object was used; this object had both a salient internal feature (an opaque white structure located inside the toy and visible through the translucent exterior of the toy) and a salient external feature (an opaque green component attached to the outer surface of the toy). Infants in both conditions were then shown two new objects. One object had the same internal white feature as the initial toy but lacked its external green feature, while the other object had the same external green feature and lacked the internal feature. Our question of interest was which toy infants would select. Infants, as stated, found the movement of the toy highly intriguing and attractive. We reasoned that they would therefore select the new toy that they expected to be most likely to make the same movement. If infants attribute self-generated motion capacities to internal or intrinsic features of an object, then infants in the Self-Generated Motion condition should be more likely to select the object with the internal feature. In contrast, infants in the Externally Caused Motion condition should choose randomly between the two objects, as neither feature is causally related to the object’s motion capacities. Infants’ choice patterns confirmed these predictions. When the object moved autonomously, infants were significantly more likely to select the object with the same internal feature. However, when the toy moved in just the same manner but through external means, they chose randomly between the two objects. By 14 months of age, infants prioritize intrinsic features over extrinsic features when reasoning about an object’s capacity for an action, but only when that action is self-generated.

2 The Nature of Infants’ Attributions to Intentional Agents

Infants have many expectations about the actions of human agents. Specifically, as discussed above, they appear to recognize that humans act in ways that are consistent with their goals or intentions. Woodward and colleagues have demonstrated that infants as young as six months old, and even three months in some circumstances, expect human hands, but not nonhuman objects, to move in goal-directed ways (Woodward, 1998; Woodward et al., 2001; Somerville, Woodward, and Needham, 2005). Ten-month-old infants expect that the “boundaries” in an action stream are related to the initiation and completion of intentions—they look longer at events in which action is paused in the middle of a goal action than at events in which the pause occurs after a goal has been completed (Baldwin et al., 2001). And 18-month-olds shown an action that an adult attempts but fails (such as trying to hang a loop on a metal prong) typically imitate the intended action even though they have never directly witnessed it (e.g., Meltzoff, 1995a). Importantly, the same infants fail to imitate the intended action when they observe the same physical motions done by a nonhuman, inanimate, mechanical hand-like device.
In addition to expecting humans to act in goal-directed ways, infants also expect that human agents will behave according to their perceptions—specifically, based on where they are looking. Sometime between 6 and 12 months, infants develop the capacity to spontaneously follow the gaze direction of an adult (e.g., Hood et al., 1998; G. Butterworth, 2001), and, soon after their first birthday, they seem to know that gaze is directed at particular objects (e.g., Brooks and Meltzoff, 2002). Infants of this age also seem to recognize that adults are more likely to perform actions on objects they are looking at. Phillips et al. (2002) found that 14-month-olds look longer when an experimenter acts on an object that she had not previously looked at, as opposed to the one that she was looking at. Finally, infants of this age also use information about where adults are looking when inferring the referent of a novel word (Baldwin, 1991, 1993) or the meaning of a negative emotional expression (Baldwin and Moses, 1994; Moses et al., 2001).

All of the examples above involve human agents, but infants seem to expect non-human agents to act in accord with their goals and perceptions as well. Infants appear capable of attributing perceptions to nonhuman agents; S. C. Johnson et al. (1998), for example, found that 12-month-olds will look in the direction in which a faceless, contingently reacting furry blob orients itself (that is, they will follow its “gaze”), consistent with the view that infants expect self-propelled contingent objects to look at interesting objects, just as a human agent does (e.g., Brooks and Meltzoff, 2002). Using computer-animated shapes moving in nonrigid, self-propelled, and contingently interactive ways, Premack and Premack (1997) found that 12-month-olds can attribute positive or negative value to the interactions of goal-directed objects: Caressing and helping actions were interpreted as having the same value, which was different from the value of hitting and hindering (which held the same value to each other). Gergely, Csibra, and colleagues (e.g., Csibra et al., 1999, 2003; Gergely et al., 1995; Gergely and Csibra, 1997) found that both 9- and 12-month-olds can develop expectations about the goal states of nonhuman agents based on their apparent goal-directedness.

However, there are different ways to explain these results. The explanation favored by some researchers (e.g., Gergely et al., 1995; Csibra et al., 1999) is in terms of infants’ understanding of what can be called “action tendencies”: Infants develop expectations about the idiosyncratic tendencies of individual agents to act toward specific goal states (“this agent tends to act toward such-and-so end state and will again do so in a future context in which this end state could obtain”), but do understand the goal state of the actor as a mental state articulating an intention (“this agent desires this or that end state and will act in accordance with this desire in the future”). Just as we all understand that all balls have a tendency to roll downhill, so infants may be able to grasp that a particular individual object (such as one of the characters in Gergely and colleagues’ experiments) has a tendency to act toward a specific goal state. Just as adults do not attribute to a rolling ball a desire to roll downhill, so infants need not be attributing a mental attitude in representing the goal action of an agent.

An alternative account posits continuity in infants’ and adults’ understanding of goals. That is, infants, like adults, are attributing genuine intentions to intentional agents—representing that this character wants to be with that character, or wishes to get to such-and-such a location or to effect a given state of affairs. To distinguish between these two possibilities, we conducted experiments in which we habituated 9- and 12-month-old infants to interactions involving a cast of characters, and then
presented the infants with test events in which the same characters appeared, but in a completely new physical context in which the old goals and tendencies did not apply because the original end state was not realizable (Kuhlmeier, Wynn, and Bloom, 2003; Kuhlmeier, Wynn, and Bloom, under review). Thus, an analysis of “action tendencies” would not generate any expectations as to how the characters might act in this new context. However, a mental-state level of analysis of the habituation situation would allow infants to generate expectations about the likely interactions of these characters in subsequent situations, including the new test context.

For example, in Kuhlmeier, Wynn, and Bloom (under review), we habituated infants to animated computer movie displays (see figure 17.3a and b) in which a little red ball-shaped agent (the Climber) attempted to climb a hill, and was alternately (a) pushed up the hill (helped) by a large yellow cube-shaped agent (the Helper) and (b) pushed down the hill (hindered) by a large green triangle-shaped agent (the Hinderer). The identity of Helper and Hinderer was counterbalanced (half the infants saw the triangle as Helper).

Infants then saw test movies in which no hill was present (see figure 17.3c). On alternate test trials the Climber approached the Helper and the Hinderer. Our question was how infants would expect the little Climber to behave in the new context? If infants interpreted the habituation event in terms of the little ball having the intention of getting to the top of the hill, and in terms of one character helping the ball achieve its goal and the other character impeding the ball from achieving its goal, then they should attribute to the ball distinct attitudes toward the Helper and the Hinderer: The ball should like, or be positively disposed toward, the Helper, and should dislike, or be negatively disposed toward, the Hinderer. The ball’s approaching the Helper is consistent with these attributions, while the ball’s approaching the Hinderer is inconsistent with them. Thus, if infants are attributing to the ball distinct attitudes to the two other characters based on its individual histories with them, infants should respond differently to events in which the ball approaches the Hinderer and events in which it approaches the Helper.

This is in fact what we found in Kuhlmeier, Wynn, and Bloom (under review).² Infants of both 9 and 12 months looked significantly longer when the ball approached

---

² In Kuhlmeier, Wynn, and Bloom (2003), we obtained a different pattern of results: Nine-month-olds showed no discrimination between the two kinds of test trials; and 12-month-olds discriminated them, showing longer looking on trials when the Climber approached the Helper. The critical difference between the events in Kuhlmeier, Wynn, and Bloom (2003 and under review) was the presence of faces on the characters in the latter, but not the former. Infants tend to prefer coherence or familiarity when faced with highly complex or difficult-to-process stimuli, and to prefer unexpected, incoherent, or unfamiliar events when faced with stimuli and events that are easy for them to process (e.g., Hunter and Ames, 1988; Hunter, Ames, and Koopman, 1983). The presence of faces on the characters (a) helps identify them as social agents (S. C. Johnson et al., 1998), and (b) helps clarify their intentions (faces give the characters the ability to orient along their direction of intended motion; as seen in figure 17.3, the little Climber faced up the hill in its attempts to climb it, thus indicating that reaching the top was indeed its intended goal). We believe that without faces on our stimuli, infants found this a difficult and challenging event to parse. The addition of faces in Kuhlmeier, Wynn, and Bloom (under review) made it sufficiently straightforward that both 9- and 12-month-olds were easily able to process the events, giving longer looking to the incoherent action of the Climber approaching the Hinderer.
the Hinderer than when it approached the Helper, indicating that they were attributing genuine mental states to the ball, and that they were perceiving our events in terms of helping and hindering. Additional experiments (see Kuhlmeier, Wynn, and Bloom, 2003) ruled out the possibility that infants merely preferred seeing the ball with one agent over the other; when test trials depicted the Helper or Hinderer approaching the ball, rather than the other way around, infants exhibited no preference for one or the other event. Their preference obtained only when the resulting pairing of ball-with-Helper/ball-with-Hinderer resulted from the ball's actions specifically. Their preferences were therefore clearly based on their expectations about the ball's attitudes and actions.

**Figure 17.3** Helping (column a) and hindering (column b) habituation events in Kuhlmeier, Wynn, and Bloom (under review).
To infer that one entity holds a particular attitude—a positive or negative inclination—toward another entity on the basis of its history with that other entity, cannot be accounted for without appeal to mental state ascriptions. There would be no call for the ball to tend toward approaching that which helped it achieve its previous goal (or to tend away from that which hindered its achievement of its previous goal) if the ball had not desired to achieve its goal.

3 Evaluating Social Actions

A prerequisite for moral thought is to be able to judge some acts as positive (or “good”), and others as negative (or “bad”). The results of the Kuhlmeier, Wynn, and Bloom (2003; under review) studies described above suggest that infants interpreted the helping act as one of positive valence, or “good,” and the hindering act as one of negative valence, or “bad.” This raises the intriguing question of whether infants in our studies were evaluating the acts, and perhaps the actors as well, in moral terms.

There has been increasing focus in evolutionary theory and comparative psychology on the possible role of moral notions in long-term adaptive behaviors such as cooperation and exchange (e.g., Axelrod, 1984; Trivers, 1971, 1985); and there have been strong claims about the extent to which human moral notions are hardwired versus the product of culture (e.g., Pinker, 1997; Turiel, 1998; Shweder, 2003; de Waal, 1996). To date, most experimental research on the origins of our understanding of moral and immoral actions has focused on children’s appreciation of justice (e.g., Damon, 1998) and the contrast between moral and conventional violations (e.g., Turiel, 1998)—and this work has been done with considerably older children. There are some findings suggesting that infants view the actions of agents in terms of “good” and “bad.” We know, for example, that infants and toddlers are distressed by the pain and distress of others (e.g., Hay et al., 1981; G. B. Martin and Clark, 1982; Sagi and Hoffman, 1976; Zahn-Waxler et al., 1992). Moreover, soon after a child’s first birthday, he or she can respond empathetically to the pain of others, trying to soothe a hurt friend, for instance (see Eisenberg and Fabes, 1991; Hoffman, 2000; Kagan, 1981, 1984; see Draghi-Lorenz et al., 2001, for even earlier estimates of empathy). More telling, there are also early signs of empathetic anger upon witnessing the harming of another child (Hoffman, 2000), as well as evidence that toddlers might
feel guilt for hurting another (Zahn-Waxler and Robinson 1995). But there is no
evidence as yet about the early development of moral attributions. Do infants, for
example, see some actions as inherently good or right, and others as bad or wrong?
While no studies yet conducted can definitively answer these questions, there is
recent work that is relevant; it addresses the somewhat broader topic of social evaluation
processes in infancy—the nature and development of the evaluative judgments
infants make about the social actions of agents.

One important study by Premack and Premack (1997) suggests that infants are
sensitive to the valence—good versus bad—of different actions. They employed a
design in which 12-month-old infants were habituated to an interaction between two
agents that was either of a positive (i.e., caressing or helping) or of a negative (i.e.,
hitting or hindering) valence. Following habituation, infants were shown two new
events, one of a positive valence and the other of a negative valence. They found
that infants looked longer at events of the new valence, suggesting that they were
encoding the events in terms of their valence.

The results of the Kuhlmeier, Wynn, and Bloom (2003; under review) studies
described above suggest that infants interpreted the helping act as one of positive
valence, or “good,” and the hindering act as one of negative valence, or “bad.” Our
next experiments investigated the nature of these evaluations in more detail. Are infants
evaluating the agents performing helping and hindering actions in positive and nega-
tive terms? It is possible that infants in our experiments described above were drawing
very “local” assessments restricted to the specific events (e.g., “the square individual
performed an act that was helpful/nice to the ball; that triangle performed an act that
was unhelpful/mean to the ball”). Alternatively, infants may have been drawing evalu-
ative inferences more broadly about the inherent nature of the agents (e.g., “that one
is a helpful/nice individual; that other one is an unhelpful/mean individual”). If the
latter is the case, then infants should not only have expectations as to the Climber’s
distinct attitudes toward the Helper and the Hinderer: Babies themselves should hold
a more positive attitude toward the Helper than the Hinderer.

To address this question, we showed both six- and ten-month-old infants the
same helping and hindering events of Kuhlmeier et al. (under review), but using
an actual (not computer-animated) display with three-dimensional material enti-
ties, with faces on them, as the agents (Hamlin, Wynn, and Bloom, under review).
Infants alternately (in counterbalanced order) saw a yellow triangular character with
eyes and a nose attempt to climb a hill and being (a) helped up the hill by a second
character, and (b) pushed down the hill by a third character.

Infants were subsequently presented with two tasks, in counterbalanced order
across infants. One was the same looking-time task as in Kuhlmeier, Wynn, and
Bloom (2003; under review): Infants saw the Climber approach, on alternate trials,
the Helper and the Hinderer, and their looking time was measured. If this three-
dimensional rendition of our helping and hindering tasks is interpreted in the same
manner as the computer-animated version used previously, then the ten-month-olds
should discriminate between the two kinds of test trials in their looking times. (It is
an open question how the six-month-olds should respond, as this age was not tested
in Kuhlmeier, Wynn, and Bloom, under review.) The other task was a choice task:
Infants were presented with a plate on which sat the Helper and Hinderer charac-
ters, and were encouraged to reach out and take one of them. If infants both (a) interpret our events as helping (positive) and hindering (negative) events, and (b) make broader inferences about the actors themselves, not just narrow evaluations restricted to the actions, then we would expect infants to reliably choose the Helper over the Hinderer as a more positive object with which to interact.

Our findings strongly supported this prediction: In our choice task, infants overwhelmingly chose the Helper rather than the Hinderer, at both six and ten months of age. This further supports the interpretation that infants are interpreting our events similarly to adults, in terms of (positive) helping and (negative) hindering actions and, moreover, suggests that infants are attributing enduring traits to the actors (“This agent is nice/good” and/or “this agent is mean/bad”) that extend beyond the specific context/situation in which the helping and hindering events take place.

It is interesting to note infants’ performance on the looking-time task. This task taps a different understanding from the choice task: It asks what expectations infants have of the Climber—what attitudes (if any) they attribute to the Climber as a result of its interactions with the Helper and the Hinderer. The choice task, in contrast, asks what attitudes the infants themselves hold toward the Helper and the Hinderer as a result of observing their respective actions toward the Climber. While the 10-month-olds performed as did the 9- and 12-month-olds in Kuhlmeier, Wynn, and Bloom (under review), showing longer looking when the Climber approached the Hinderer than when it approached the Helper, the six-month-olds showed no looking time preference for either event. Our finding that six-month-olds systematically choose the Helper in the choice task, but do not discriminate the two actions of the Climber in the looking-time task, suggests that evaluating social actions and generating one’s own attitudes toward an actor occur earlier in ontogeny than does the attribution of attitudes to others.

Infants’ choices in our above experiment could be completely accounted for solely by a tendency to be drawn toward the Helper, or solely by a tendency to avoid the Hinderer; alternatively, both processes could be at work. To ask separately (a) if infants positively evaluate an individual who helps another in attaining a goal, and (b) if infants negatively evaluate an individual who impedes another in attaining a goal, we conducted two further experiments. In our “Is a Helper Good?” experiment, six- and ten-month-old infants were given a choice between a Helper and a Neutral character. Infants were first habituated to two alternating events. In one event, an individual, the “Helper,” helped a second individual (the “Climber”) up a hill. In the other event, a third individual moved up the hill, tracing the same path and sequence of motions as the Helper, but without contacting (and thus without in any way interacting with or helping) the Climber. This third individual we called our Neutral character. Similarly, in our “Is a Hinderer Bad?” experiment, six- and ten-month-olds chose between a Hinderer and a Neutral character. Infants were habituated to a Hinderer who pushed the Climber down the hill it was attempting to climb, and to a Neutral character who traced the path of the Hinderer, and in the same manner, but without contacting or interacting with the Climber.

If infants are making positive judgments about Helpers, then in the first of these experiments they should choose the Helper over the Neutral character. If they are making negative judgments about Hinderers, then in the second experiment they
should choose the Neutral character over the Hinderer. Indeed, for both six- and ten-month-old infants, this was the pattern of results we obtained: The strong majority of infants of both ages selected the more positively valenced character in each experiment. That is, even by six months of age, infants both find helping characters attractive, and find hindering characters aversive, indicating that within the first half-year of life infants are making both positive and negative evaluations of individuals, based on the nature of their social actions.

I do not claim that these evaluations comprise a system of moral cognition in infants. There are many components to a system of moral judgment, including the capacity to weigh the multiple and conflicting needs and rights of many different actors; the ability to consider and prioritize abstract principles of fairness and justice, considerations of punishment, reward, mercy, and forgiveness; and numerous other components that are surely far beyond the conceptual grasp of preverbal infants. I do suggest that these early social evaluations provide an essential foundation upon which a developing system of moral cognition can be built. These judgments possess at least one of the important properties of genuine moral judgments (in contrast to merely valence judgments): They are abstract judgments in the sense that they are unconnected to the infant’s own personal experience and welfare. The infants in our studies have had no personal history with the agents used in our experiments, nor have they themselves experienced any (positive or negative) consequences as a result of the agents’ actions. Their evaluations are made on the basis of witnessed interactions between unfamiliar individuals; the infant, as an unaffected, unrelated—and therefore unbiased—third party, is nonetheless rendering an abstract judgment about the value of a social act.

These studies are also relevant to the question of whether infants make genuine mental attributions from an early age. Even 6-month-old infants—significantly younger than the 9- and 12-month-olds in the studies of Kuhlmeier, Wynn, and Bloom reviewed in section 2 above—judged some social actions as positive and others as negative. What gives (e.g.) a hindering action its negative valence? The only answer can be that it is having not just an effect on the acted-upon agent, but a negative effect on that agent. But impeding the tendency of a nonintentional object cannot generally be viewed as negatively affecting that object, though it certainly does affect the object in question. It is not a negative action to stop a stone’s rolling downhill (though smashing the stone could possibly be viewed as a negative action). What makes the Hinderer’s action in our studies a negative one is the thwarting of an intention, not the preventing of a tendency; the foiling of the Climber’s desire. This suggests that in order for our six-month-olds to judge the Hinderer negatively, they had to be understanding the goal of the Climber in intentional terms. A similar analysis, of course, holds for infants’ positive evaluation of a helping act—such an act is positive only by virtue of its relationship to the intentions and desires of the helped object.

4 Conclusions

I have reviewed research showing that young infants distinguish the social world from the physical world. They have distinct, possibly even nonoverlapping, systems for reasoning about each. While infants reason about inanimate objects in terms
of naïve physics, they reason about the behavior of social beings through appeal to mental states such as goals, desires, and attitudes. These systems are sufficiently distinct that under at least some conditions, infants do not even think of social beings as material objects. Moreover, infants make different inferences about the source of behavior of intentional agents as opposed to inanimate objects. They view actions of an intentional agent as arising from intrinsic, essential features of the agent, not its superficial or extrinsic attributes. Infants also make inherently mentalistic attributions to social beings in making sense of their behavior, and they assign value to their actions—and to the actors themselves—from the earliest ages.

These claims collectively paint a picture in which the infant’s initial impression is of a material domain comprised of physical objects constrained by a naïve physics, and of a social domain comprised of intentional beings with mental states and content. But these beings are not merely a sum of isolated individuals, each with a psychology; rather, they are inextricably part of a community of actors within which an individual’s acts gain their meaning by virtue of their social causes and consequences. In short, the social world is not one the infant must laboriously create; it is one that is presupposed.